

DOCUMENT RESUME

ED 476 416

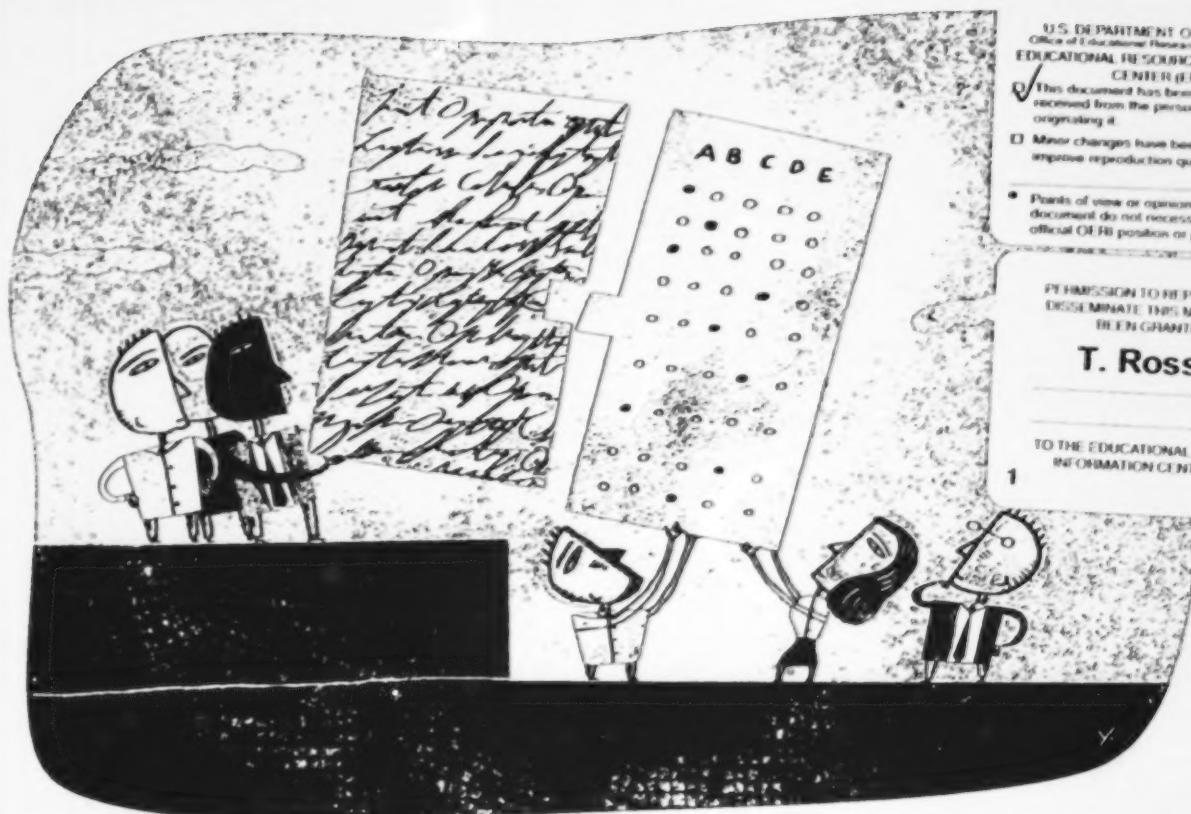
TM 034 915

AUTHOR Ananda, Sri
TITLE Rethinking Issues of Alignment under "No Child Left Behind."
Knowledge Brief.
INSTITUTION WestEd, San Francisco, CA.
SPONS AGENCY Institute of Education Sciences (ED), Washington, DC.
PUB DATE 2003-00-00
NOTE 13p.
CONTRACT ED-01-CO-0012
PUB TYPE Reports - Descriptive (141)
EDRS PRICE EDRS Price MF01/PC01 Plus Postage.
DESCRIPTORS *Academic Achievement; Curriculum Development; Elementary
Secondary Education; *Research Methodology; *Standards
IDENTIFIERS *Curriculum Alignment; *No Child Left Behind Act 2001

ABSTRACT

This Knowledge Brief is intended for designed and consumers of alignment studies as they begin or continue to address this issue. It describes different ways of conceptualizing and examining alignment, especially those brought to the fore by the No Child Left Behind Act. The brief first defines alignment and compares it to related concepts. Alignment refers to the degree to which standards, assessments, and other important elements in an education system are complementary and work together to gauge student learning effectively. The brief then discusses key alignment issues that states and districts now face, providing examples of the various purposes served by alignment studies and different ways to achieve alignment. It also touches on methodological advances in the study of alignment. The brief concludes with recommendations for states and districts intending to conduct or commission alignment studies. (Contains 13 references.) (SLD)

ED 476 416



U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

✓ This document has been reproduced as
received from the person or organization
originating it.

□ Minor changes have been made to
improve reproduction quality.

• Points of view or opinions stated in this
document do not necessarily represent
official OEI position or policy.

PERMISSION TO REPRODUCE AND
DISSEMINATE THIS MATERIAL HAS
BEEN GRANTED BY

T. Ross

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)

1

**Rethinking
Issues of**

Alignment

Under No Child Left Behind

That standards and assessments must be properly aligned is neither new nor controversial. But the need for alignment has acquired new urgency with the escalating use of student assessment results to determine sanctions and rewards for schools, teachers, and students. This trend is embodied most visibly in the No Child Left Behind Act of 2001 (NCLB). Its myriad assessment-and-accountability provisions are pushing states and districts to address many alignment issues that have often received short shrift.

Written by Sri Ananda

WestEd

*Improving education through
research, development, and service*

This Knowledge Brief is intended for designers and consumers of alignment studies as they begin or continue to address this important issue. It describes different ways of conceptualizing and examining alignment — particularly those brought to the fore by NCLB. The brief first defines alignment and compares it to related concepts. It then discusses key

TM034915

alignment issues that states and districts now face, providing examples of the various purposes served by alignment studies and different ways to achieve alignment. It also touches on methodological advances in the study of alignment. The brief concludes with recommendations for states and districts intending to conduct or commission alignment studies.

What is alignment?

Simply stated, alignment means agreement. In broad terms, it refers to the degree to which standards, assessments, and other important elements in an education system are complementary and work together to effectively guide student learning (Webb, 1997). The underlying assumption is that a coherent message and system will positively influence what teachers teach and what students ultimately learn. As such, alignment is considered pivotal to a standards-based reform system (Porter, 2002; Smith & O'Day, 1991).

Most often, those considering alignment focus specifically on the degree of match between test content and expected learning content (e.g., content standards). Along with other aspects of alignment, NCLB also focuses on this essential piece. Although different researchers offer somewhat different criteria for alignment of standards and assessment, all agree that, minimally, alignment requires content match and depth match (La Marca, 2001). Content match refers to how well test content corresponds to targeted standards. Depth match refers to how well test items reflect the cognitive complexity of the knowledge and skills specified in the standards. An example of weak or superficial depth match is when a given math item asks students to identify the correct geometric shape (e.g., "Which of the following is a trapezoid?"), but the

content standard states that students should know how to "calculate the area or perimeter of triangles and quadrilaterals." Even though both the assessment item and standard cover the same general skill category (knowledge of geometric shapes), the item addresses the skill at a lower complexity level than represented by the standard.

The concept of alignment is intricately connected to test validity, the notion that accurate inferences can be drawn from test scores. Specifically, alignment is closely allied with at least three different aspects of test validity. First, it relates to content validity, the degree to which test content is representative of the targeted content domain. Evidence of strong alignment between test content and standards is typically used as evidence of content validity.

**MINIMALLY, ALIGNMENT
REQUIRES CONTENT MATCH
AND DEPTH MATCH.**

Alignment is also closely connected to construct validity, "the evidential basis of test interpretation" (Messick, 1989, p. 34). Construct validation

investigates evidence of underlying abilities or traits (i.e., constructs) to account for and explain student test performance. Alignment studies can provide evidence of construct validity. For example, "reading with understanding" is a developmental construct that is hypothesized to change over time as children develop as readers. If an alignment study were to find that a state's or district's comprehension-related reading standards and assessments were aligned and articulated across grade levels in a way that reflected these presumed changes, that finding would be considered evidence of construct validity. Finally, alignment in its broadest sense relates to evidence of consequential validity. Both are ultimately concerned with the social consequences of testing, including the desired outcome of improved student learning.

What does NCLB say and suggest about alignment?

First and foremost, NCLB explicitly says that states may select and design assessments of their own choosing, with the stipulation that state assessments be aligned with state academic standards. NCLB also calls for grade-specific content standards and annual testing for grades 3 through 8, as well as reporting of assessment results by standards. These requirements suggest the need for meaningful alignment, or articulation, of standards and assessments across grade levels.

For states electing to use a norm-referenced test (NRT), the NCLB regulations specify that they must perform an independent alignment study relative to the state's content standards to identify missing content and then augment the assessment accordingly. Some states are now considering whether to omit from their augmented assessments any NRT items that do not align to their state content standards. Among those states that have decided to keep the nonaligned NRT items, some are planning to report two separate scores for NRT results: (1) a score based on all NRT items, which would indicate student achievement relative to established norms and (2) a score based only on the aligned items, which would indicate student achievement relative to content standards.¹

As is true of its assessment requirements, NCLB's accountability provisions also raise interesting alignment issues. Briefly, NCLB requires states to set up a single, statewide accountability system in which, by the end of 12 years, all students will be expected to score at a state-defined "proficient" level. Furthermore, in moving toward this goal of 100 percent proficiency, each state must show that there is a gradual improvement in student performance

occurring in equal increments over time. To comply with this mandate, many states are considering how to establish appropriate passing scores for different grade levels on state or local tests. Doing so requires serious attention to alignment. Misalignment between standards across different grades or between standards and assessments within particular grades could result in a state setting inappropriate passing scores, leading to variable proficiency rates for students at different grade levels. Having widely different proficiency rates for students at different grade levels would be difficult for a state or district to explain or defend.



Another significant alignment issue is raised by NCLB's required use of the National Assessment of Educational Progress (NAEP) as an external accountability indicator, that is, as a point of comparison for a state's results on its own tests. Beginning in 2002/03, a sample of 4th and 8th graders in each state must participate in NAEP reading and math tests every other year. Although states will not be sanctioned for low NAEP performance, such results could raise questions about the quality of a state's testing program, particularly if the NAEP results differ significantly from the state test results. Interestingly, an analysis by *Education Week* in early 2002 focusing on the 25 states that participated in NAEP mathematics tests in 2000 showed that states varied widely with respect to how closely their state assessment achievement levels relate to their NAEP achievement levels. A recent alignment study by WestEd (2002) demonstrated that four states showed considerable content overlap between their content standards and NAEP frameworks, but with notable differences in content emphases and depth of content coverage. With NCLB's new use of NAEP, states must decide the extent to which they want their assessments and, therefore, academic standards to align with NAEP.

Finally, more and more states now realize that in order to have an effective overall education reform system, their assessment and accountability components must be aligned. Implementation of NCLB raises significant challenges for states seeking this alignment (Carlson, 2002). For example, NCLB presents conflicting timelines for assessment and accountability requirements. The law mandates that states establish their accountability (i.e., adequate yearly progress) baseline measurements in 2001/02, yet it does not require that any new assessments begin implementation until 2005/06. This means states must establish baseline measurements using existing assessments, then switch to the new measures midstream as they pursue their goal of 100 percent student proficiency over 12 years. In light of this switch, the consistency and continuity of accountability decisions over time (e.g., which schools get identified as successful or in need of improvement) may be difficult to maintain and defend.

UNDER NCLB, STATES ELECTING TO USE A NORM-REFERENCED TEST MUST PERFORM AN INDEPENDENT STUDY TO IDENTIFY MISSING CONTENT AND THEN AUGMENT THE ASSESSMENT ACCORDINGLY.

NCLB also presents content-related challenges for alignment of assessment and accountability system components. For example, in the interest of maximal inclusion of students, the law's assessment provisions allow some students with special needs to be assessed with accommodations specified in students' individualized education plans (IEPs). Yet the law's accountability provisions state that irrespective of IEPs, special needs students who have been tested with some accommodations in test format and content cannot be counted as meeting the proficiency standard of the accountability system. This means that while a school may meet NCLB's assessment requirements by testing special needs students with IEP-specified accommodations, the same school may fall short of its accountability goals (i.e., adequate yearly progress)

because the use of these accommodations automatically means a student cannot be considered "proficient" for accountability purposes.

What types of alignment studies should states be asking for?

In this era of high-stakes testing and accountability, states and districts are asking for alignment studies to serve a wide range of purposes. Purposes for alignment studies range from exploratory to confirmatory, and from less formal to more formal. Among the purposes alignment studies can serve are:

Identify areas of vulnerability. Faced with the ambitious assessment-and-accountability implementation timelines associated with NCLB and other federal or state mandates, several states are seeking targeted, short-term alignment studies. These studies are geared to quickly providing states

with answers to questions such as: Where are we vulnerable? Where are our content gaps and other potential shortcomings? For example, in light of NCLB, states are now asking for alignment studies to help determine whether their existing tests sufficiently represent the breadth and depth of their state content standards. Many of these states are willing to forego complex alignment research designs and detailed final reports to obtain fairly quick and global information about their content gaps so that they can quickly begin to strategize next steps.

Inform restructuring of an existing assessment or accountability system. Whereas some alignment studies are intended only to signal possible problems, others are designed to provide more specific and

detailed information to help states and districts decide whether to restructure their existing systems, or how to go about restructuring their systems. Evidence of strong alignment might argue for slight changes to an existing system, whereas evidence of weak alignment might argue for major restructuring or even building a new system from scratch.

Guide vertical scaling. Scaling is the process of transforming raw test scores (e.g., 30 correct answers out of 40 questions) into a format that allows for clearer interpretation and comparisons of student achievement. For example, placing student scores from different forms of a grade 4 science test on a scale with a pre-established mean and standard deviation allows the comparison of scores across equated forms of that test. Vertical scaling goes one step further, placing scores from the same test at different grades (e.g., science tests at grades 3 through 8) on the same scale. To facilitate comparisons and interpretation of growth in student achievement, NCLB encourages states to use formal vertical scales across the grades that will be tested.

However, use of vertical scales assumes proper alignment of standards across the grade levels to be tested. Thus, districts and states that contemplate using a vertical scale should conduct an alignment study examining the articulation of standards across grade levels to ensure the validity of a vertical scale.

Compare own standards and assessments to others.

States and districts sometimes want to know how their standards and assessments compare to those in other systems, building such comparisons into their alignment studies. If comparisons are desired, states or districts must decide on the most appropriate basis for comparison. In some instances, they want to know how

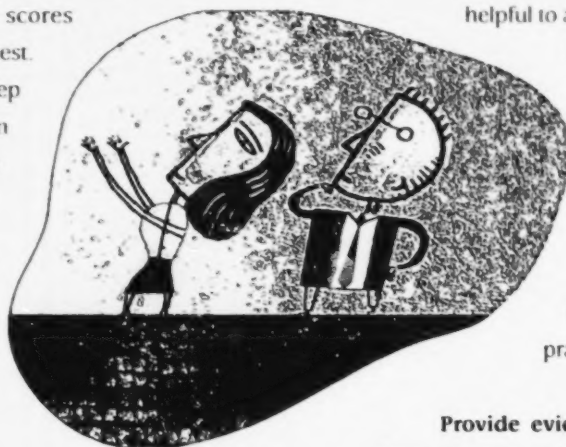
they compare to standards and assessments that are known to be exemplary. In other instances, states want to know how their systems compare to those of other states in the same region or serving similar student populations. Alternatively, some states have explicitly asked that the alignment studies they commission not include comparisons to others.

Inform future assessment item development activities

Alignment studies can serve the very utilitarian purpose of guiding the augmentation of assessment item pools. When designed appropriately, an alignment study might identify gaps or surpluses in item pools relative to targeted standards, as specified by the state's assessment plans or blueprints. Such information can be extremely

helpful to a state or district in planning

its future item development activities. Although future item development is not typically offered to the public as the primary rationale for an alignment study, it is nonetheless an important practical consideration.



Provide evidence of content validity

from an external source. Sometimes states

and districts want to obtain confirmatory evidence of the content validity of their assessments from an independent, expert source (e.g., someone other than their own staff or development contractors). They may seek this evidence for their own purposes or because it has been requested by others, such as the state legislature or the board of education. In either case, states and districts would then commission an external study to investigate the alignment of their standards and assessments. When commissioning an external alignment study, states and districts must be explicit about the specific alignment questions that interest them and the desired features of the study.

Alignment Study Purposes and Their Implications for Study Design

Purpose of Alignment Study	Implications for Study Design
Identify areas of potential vulnerability and shortcomings in standards or assessments	A targeted, short-term alignment study may suffice if the goal is simply to signal potential problems.
Inform restructuring of an existing assessment or accountability system	An alignment study must yield comprehensive, detailed information if it is to help a state or district decide how to restructure its existing system.
Guide vertical scaling	A district or state contemplating use of a vertical scale to report assessment results must first conduct an alignment study that examines the articulation of standards across grade levels to ensure validity of the vertical scale.
Compare own standards and assessments to others	A state or district must first determine what would be the most relevant comparisons (e.g., to other districts or states that share key student demographic characteristics; to other standards and assessments that have been deemed exemplary).
Inform future assessment item development activities	Although it is not typically the primary rationale for an alignment study, states and districts may want to incorporate this practical goal into an alignment study, regardless of the study's primary purpose.
Demonstrate compliance with a mandate	An alignment study undertaken specifically to demonstrate compliance with a mandate should be conducted by an external party so as to be perceived as more objective.

Demonstrate compliance with a mandate. Irrespective of any other purposes for an alignment study, a common reason for commissioning one is to show compliance with a district, state, or federal mandate. The advantage of having such studies conducted by an outside party is that their evidence of compliance will be perceived as more objective. As previously described, NCLB requires that states using NRTs commission an external alignment study and then augment the NRT as needed to ensure that it adequately aligns with state standards. Thus, states using NRTs must conduct and document the results of such alignment studies simply to be in compliance.

How is alignment achieved?

Alignment can be achieved through use of sound standards- and assessment-development practices that focus on alignment during each step of the process. For example, during standards development, one

should proactively consider how achievement will be measured. One way is to design exemplary assessment items *during* the standards development process, thereby providing concrete examples of the desired relationship of a given standard to its assessment. These exemplary items could then be used to inform subsequent assessment development, bridging the standards- and assessment-development processes.

Similarly, attention to alignment can be built into other front-end assessment-development practices. Test blueprints and item specifications should clearly stipulate the connections of standards and objectives to assessment content and format. Schafer (in progress) advocates use of a test map, which is much more informative than typical test blueprints and item specifications. A test map explains in detail for teachers and test developers what is "fair game" to be included in the test, enumerating content and cognition limits, assessment formats, how test items will be sampled, and what subscores will be developed and reported.

As a prerequisite to item development, item writers must become intimately familiar with the targeted standards and test blueprint (or test map), designing and then coding their draft items to the appropriate standards. During the item review process, content experts should evaluate the closeness of the relationship between each assessment item and the standard(s) it is intended to measure. Items found to have weak or superficial connections with the targeted standards should then be revised or dropped.

Once an assessment is fully developed, a formal alignment analysis should be conducted. To obtain a complete picture of the relationship between standards and assessments, it is important to examine the alignment from two directions: (1) how well each targeted standard is covered by the assessment and (2) what proportion of the overall assessment is aligned to the standards. Many alignment studies focus on one direction or the other, but not both.

What is sufficient alignment between standards and assessments?

Conventional wisdom holds that sufficient alignment between standards and custom-developed criterion-referenced tests (CRTs) is nearly guaranteed, provided that the CRT is designed to be consistent with specific standards and that during the development process the designers have each item reviewed by experts in its respective content area. However, this assumption about sufficient alignment does not always stand up under scrutiny. In fact, a number of researchers now assert that states have a lot of work to do to bring their standards and CRTs into alignment (Webb, 1999). That said, there is no hard and fast rule about what constitutes sufficient

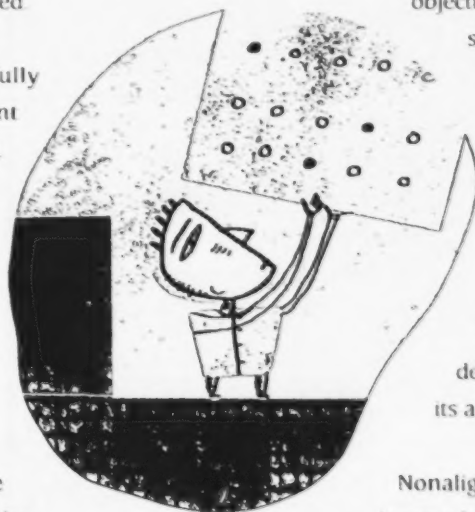
alignment. Indeed, defining sufficient alignment is complicated by a number of factors.

First, it is common for content standards to include different levels of content specificity. Some assessments are designed to reflect targeted standards at a global level (e.g., content strand, content standard), whereas others are designed to align to a more specific level of the content standard (e.g., grade-level benchmark, objective, or performance indicator). If a

state has designed its assessment to align to standards at a global level, an alignment study focusing on the relationship between standards and assessment at a more specific or fine-grained level may conclude that there is insufficient alignment. Whether this conclusion is justified would depend on the state's intention for its assessment.

Nonalignment may also come up as an issue when one or more standards are not amenable to measurement. For example, "students develop appreciation for reading" may be a worthwhile standard around which to design a curriculum, but it is also a standard that would be difficult to measure. Similarly, some standards that call for students to produce extensive work products or projects (e.g., laboratory experiments, case studies) may not easily be accommodated in large-scale assessment, but may be appropriate for performance-based classroom assessment.

Weak alignment of standards and assessments can also result from flaws in either the standards or assessments, or both. Specifically, many alignment studies have uncovered deficiencies in standards such as lack of clarity in particular standards or redundancy of content across different standards. Such deficiencies make it difficult to develop items that are sufficiently aligned. Similarly, it is often difficult to evaluate the



match of ill-constructed assessment items to particular content standards.

Another complication in determining whether alignment is sufficient comes up when tests are intentionally designed to align to standards at more than one grade level. For example, high school exit exams are intended for high school students, yet many of these exams include items that are targeted to both high-school-level standards and standards for lower grade levels (e.g., eighth grade computation). Sometimes states deliberately opt for nonalignment to grade-level standards when field test data show that large numbers of students would otherwise fail the high school exit exam. In these cases, a decision is made to remove harder items (e.g., algebra, geometry), even though the result is a test less aligned to grade-level content standards.

Finally, as Porter (2002) points out, tests comprise a sample of items from a domain, whereas standards make up the domain. "Thus, perfect alignment should not be expected" (p. 6). In defining the content domain, many state content standards in use today were developed primarily to guide and evaluate instruction at the classroom level, not to serve as the basis for evaluating an individual student's learning. Most sets of content standards cover far too much content than any given student could be expected to master. The comprehensiveness of standards limits the extent to which standards and assessments can be aligned.

Given these complicating factors and the general lack of agreement about what constitutes sufficient alignment, determining whether a state's or district's system is sufficiently aligned is somewhat arbitrary. Thus, to make any alignment study more meaningful,

states and districts must select among different options available to them for demonstrating alignment (e.g., use of CRTs vs. augmented NRTs; selecting among different alignment methodologies). Pursuing fuller alignment involves strategic trade-offs that may require more time (e.g., increased test length) and resources (e.g., money for developing and scoring open-response items) than are readily available to a state or district.

What methods are used to examine alignment?

Methodologies for investigating alignment are still in their youth. In the past, many alignment studies, using different, and sometimes unclear, methodologies, were criticized as being arbitrary and subjective. However,

methodologies that show great promise are now emerging.² They provide more detailed, in-depth measures of alignment between standards, assessments, and instruction, pinpointing where "standards and assessments intersect and where they do not" (Council of Chief State School Officers, 2002, p. 1). Most entail a systematic review of the standards by content experts, followed by their systematic review of assessment items and tasks.

These experts are trained to judge alignment against a specific set of alignment criteria and decision rules. More than one reviewer is used to judge alignment between any given set of standards and assessments, and agreement among reviewers (i.e., inter-rater reliability) is monitored.

Webb (1997, 1999), for example, has written extensively about an alignment analysis process that calls for content experts to use the following alignment criteria:

**THE COMPREHENSIVENESS
OF STANDARDS LIMITS
THE EXTENT TO WHICH
STANDARDS AND
ASSESSMENTS CAN
BE ALIGNED.**

Categorical Concurrence: How similar are the categories of content in the standards and assessments?

Depth-of-Knowledge Consistency: To what extent is the knowledge tapped by the assessment as cognitively demanding as that reflected in the standards?

Range-of-Knowledge Consistency: To what extent is the span of knowledge represented in the assessments comparable to that represented in the standards?

Balance of Representation: How evenly are the assessment items distributed across the various objectives within a standard?

Indeed, one of Webb's many significant contributions to the study of alignment is his model's systematic focus on different key aspects of content alignment between assessments and expectations (e.g., standards).

In fact, others are beginning to use Webb's rich description of alignment criteria in their front-end development practices to help ensure proper alignment as they develop or revise assessments (Petit, 2002).

Achieve's approach to examining alignment uses criteria that overlap with Webb's, including content centrality, performance centrality, challenge, balance, and range (Achieve, 2002). In addition to providing qualitative and quantitative analysis on the alignment of a state's assessments to its content standards, Achieve offers services such as a comparison of a state's standards to exemplary state and international standards, and a comparison of norm-referenced tests to a state's content standards.

The Council for Basic Education (CBE) also has its own alignment process, which it has used in a number of alignment studies. In this approach, reviewers work in pairs, applying a scoring rubric and exemplars to evaluate the degree of match between test items and standards (or bench marks). The four alignment criteria they focus on are content, content balance, rigor, and item response type.

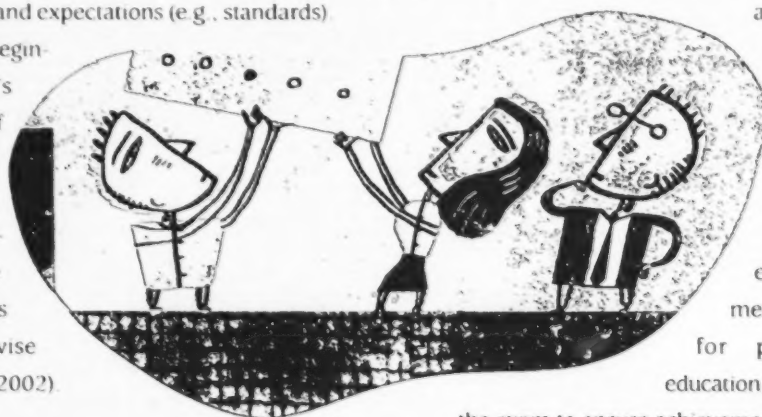
WestEd's own approach to alignment studies typically features a highly collaborative process. That is, WestEd actively engages the state or district whose standards and assessments are being studied in determining alignment criteria and decision rules. This helps ensure that the results of the study will provide the specific types of information that each particular state or district needs to improve its systems, consistent with the stated purposes of the education reform model its standards

and assessment were developed to support.

For example, some states intend their high school exit examination to measure readiness for postsecondary education while others use the exam to ensure achievement of basic skills.

Such differences in intent guide the alignment criteria used in the WestEd alignment analyses. Clearly, the inherent challenge in WestEd's approach is to maintain the independence and neutrality of an external review even as reviewers and client collaborate.

In deciding which alignment study approach to use, states and districts must consider their own particular alignment goals. For example, some states or districts may not value as a goal for their assessments the balanced distribution of assessment items across the various objectives within a standard. They may have intentionally built into their assessments greater emphasis on some



objectives within a standard than on others, avoiding those objectives deemed unsuitable for large-scale assessment.

Whereas most approaches to alignment do not allow for direct state-to-state comparisons of the alignment between standards and assessments, Porter (2002) offers innovative and promising alignment indices that are expressly used to compare alignment within and between states. These indices allow such comparisons by mapping standards and assessments from different states onto a common content language. Porter cites a recent study that used one of his alignment indices to demonstrate that the grade 7 mathematics assessment in each of four states is no more aligned to its own standards than it is to the content standards of the other states or to those of the National Council of Teachers of Mathematics. This provocative finding raises several interesting questions about the current status and nature of alignment, including whether states need to do more to align their standards and assessments or whether there needs to be further examination of what constitutes sufficient alignment. While researchers will undoubtedly want to pursue use of Porter's indices to help answer such important questions, when considering possible alignment approaches and criteria, states and districts need to determine for themselves whether comparisons to other states are relevant for their particular alignment purposes.

Conclusion

Designers and consumers of alignment studies should consider the following steps that cover both the planning of an alignment study and use of its results to inform practice:

- Be clear about the purpose(s) of the alignment study and how the results will be used.
- Select a methodology that matches your purpose and resources, remembering that not all methods lead to the same conclusions.
- Determine appropriate alignment criteria and decision rules.
- Review the results of your alignment study to determine if there is evidence of sufficient alignment.
- Secure resources and approvals necessary to implement changes to your system based on the findings of your alignment study.
- Refine your system by building alignment practices into front-end processes, and monitor alignment throughout the life of your system.

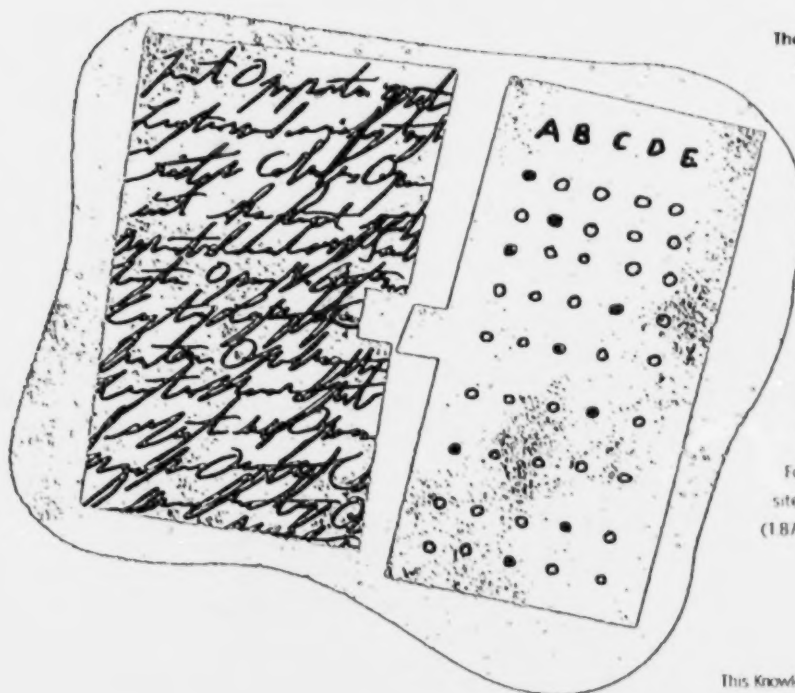
Endnotes

¹ Any changes to the content or administration of an NRT could invalidate the norm-based scores. Therefore, more research is needed to determine how robust NRT norms are given possible NCLB-inspired changes to test content and administration conditions.

² See the Council of Chief State School Officers' *Models for Alignment Analysis and Assistance to States* (2002) for a comparative analysis of alignment models.

References

- Achieve. (2002) [on-line]. *Benchmarking*. Available online: <http://www.achieve.org/achieve.nsf/Benchmarking>.
- Carlson, D. (2002). Toward a structure for evaluating and improving the validity and reliability of state accountability systems. In S. Erpenbach (Ed.), *Making valid and reliable decisions in the determination of adequate yearly progress*. Washington, DC: Council of Chief State School Officers.
- Council of Chief State School Officers. (2002). *Models for alignment analysis and assistance to states*. Available online: <http://www.ccsso/pdfs/AlignmentModels.pdf>.
- Marca, P. M. (2001). Alignment of standards and assessments as an accountability criterion. *Practical Assessment, Research & Evaluation*, 7(21). Available online: <http://ericae.net/pare/getvn.asp?v=7&n=21>.
- Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd ed.). New York: American Council on Education — Macmillan Publishing Company.
- Olson, L. (2002, February 20) A 'proficient' score depends on geography (with accompanying charts, What's proficient?). *Education Week on the Web*. Available online: <http://www.edweek.org/ew/newstory.cfm?slug=23proficienth21>.
- Petit, M. (2002, October). *Improving the No Child Left Behind Act*. Presentation at Reidy Interactive Lecture Series, Nashua, NH.
- Porter, A. C. (2002). Measuring the content of instruction: Uses in research and practice. *Educational Researcher*, 31(7), 3-14.
- Schafer, W. D. (in progress). *Designing assessments for teaching*. College Park, MD: University of Maryland.
- Smith, M. S., & O'Day, J. (1991). Systemic school reform. In S. H. Fuhrman & B. Malen (Eds.), *The politics of curriculum and testing: The 1990 yearbook of the Politics of Education Association* (pp. 233-267). Bristol, PA: Taylor & Francis.
- Webb, N. L. (1997). *Research monograph no. 6: Criteria for alignment of expectations and assessments in mathematics and science education*. Washington, DC: Council of Chief State School Officers.
- Webb, N. L. (1999). *Alignment of science and mathematics standards and assessments in four states*. Washington, DC: Council of Chief State School Officers.
- WestEd. (2002). *A comparison of NAEP content to reading, writing, and mathematics content standards for Arizona, California, Nevada, and Utah*. San Francisco: Author.



The author would like to acknowledge the helpful review of this Knowledge Brief by members of the National Assessment and Accountability Work Group.

WestEd, a nonprofit research, development, and service agency, works with education and other communities to promote excellence, achieve equity, and improve learning for children, youth, and adults. While WestEd serves the states of Arizona, California, Nevada, and Utah as one of the nation's Regional Educational Laboratories, our agency's work extends throughout the United States and abroad. It has 16 offices nationwide, from Washington and Boston to Arizona, Southern California, and its headquarters in San Francisco.

For more information about WestEd, visit our Web site: WestEd.org/, call 415 565 3000 or, toll free, (1 877) 4 WestEd, or write:

WestEd
730 Harrison Street
San Francisco, CA 94107-1242

This Knowledge Brief was produced in whole or in part with funds from the Institute of Education Sciences, U.S. Department of Education, under contract #ED-01-CO-0019. Its contents do not necessarily reflect the views or policies of the Department of Education.

© 2003 WestEd. All rights reserved.

WestEd®

730 Harrison Street
San Francisco
California 94107-1242

Address service requested

Non-Profit
U.S. Postage
P A I D
Los Alamitos, CA
90720
Permit No. 87

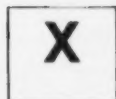


U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)



NOTICE

Reproduction Basis



This document is covered by a signed "Reproduction Release (Blanket)" form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.



This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").